Reliable and Interpretable Mortality Prediction With Strong Foresight in COVID-19 Patients: An International Study From China and Germany

Tao Bai[1†], Xue Zhu[2†], Xiang Zhou[3], Denise Grathwohl[3], Pengshuo Yang[2], Yuguo Zha[2], Yu Jin[1], Hui Chong[2], Qingyang Yu[2], Nora Isberner[3], Dongke Wang[1], Lei Zhang[1], K. Martin Kortüm[3], Jun Song[1], Leo Rasche[3], Hermann Einsele[3], Kang Ning[2]* and Xiaohua Hou[1]*

[1]Division of Gastroenterology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2]Key Laboratory of Molecular Biophysics of the Ministry of Education, Hubei Key Laboratory of Bioinformatics and Molecular-imaging, Department of Bioinformatics and Systems Biology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, China, [3]Department of Internal Medicine II, University Hospital of Würzburg, Würzburg, Germany

Cohort-independent robust mortality prediction model in patients with COVID-19 infection is not yet established. To build up a reliable, interpretable mortality prediction model with strong foresight, we have performed an international, bi-institutional study from China (Wuhan cohort, collected from January to March) and Germany (Würzburg cohort, collected from March to September). A Random Forest-based machine learning approach was applied to 1,352 patients from the Wuhan cohort, generating a mortality prediction model based on their clinical features. The results showed that five clinical features at admission, including lymphocyte (%), neutrophil count, C-reactive protein, lactate dehydrogenase, and α-hydroxybutyrate dehydrogenase, could be used for mortality prediction of COVID-19 patients with more than 91% accuracy and 99% AUC. Additionally, the time-series analysis revealed that the predictive model based on these clinical features is very robust over time when patients are in the hospital, indicating the strong association of these five clinical features with the progression of treatment as well. Moreover, for different preexisting diseases, this model also demonstrated high predictive power. Finally, the mortality prediction model has been applied to the independent Würzburg cohort, resulting in high prediction accuracy (with above 90% accuracy and 85% AUC) as well, indicating the robustness of the model in different cohorts. In summary, this study has established the mortality prediction model that allowed early classification of COVID-19 patients, not only at admission but also along the treatment timeline, not only cohort-independent but also highly interpretable. This model represents a valuable tool for triaging and optimizing the resources in COVID-19 patients.

Keywords: COVID-19, Wuhan cohort, Würzburg cohort, mortality prediction model, reliability, interpretability, foresight

INTRODUCTION

The pandemic of coronavirus disease 2019 (COVID-19) has become a public health emergency of international concern (Salyer et al., 2021; Sirleaf and Clark, 2021; Watson and Lilford, 2021). As of July 12, 2021, 187, 796, 841 confirmed infection cases have been reported by the World Health Organization, with a global mortality rate of 2.16% (https://covid19.who.int/). Even worse, the incidence of COVID-19 is continuously increasing worldwide, and areas already under control are likely to relapse (Setti et al., 2020). The proportion of critically ill COVID-19 patients is 18.5% (Epidemiology Working Group for Ncip Epidemic Response CCfDC, Prevention, 2020), and this high proportion of severe cases has put enormous pressure on medical systems, resulting in a serious shortage of medical resources (Rasmussen et al., 2020; Ammar et al., 2021; Wahlster et al., 2021).

In recent years, machine learning methods used for large clinical data analysis have been sprung up (Liang et al., 2020; Wu et al., 2020; Xiao et al., 2020; Zhu et al., 2020; Gomes and Serra, 2021; Ikemura et al., 2021; Wang et al., 2021). Yan et al. used the XGBoost classifier (Chen and Guestrin, 2016) to predict the outcome of 485 patients using the final samples at discharge, and they found three blood features that could be used as predictors, providing important evidence for clinical decision-making and patient management (Liang et al., 2020). Xiao et al. have used the HNC-LL score that considered hypertension, neutrophil count, C-reactive protein (CRP), lymphocyte count, and lactate dehydrogenase (LDH) to predict the severity of COVID-19 with AUC higher than 0.82 based on 442 patients (Xiao et al., 2020). Liang et al. developed a deep learning survival Cox model for 1,590 patients' triage, which was based on four clinical features and six phenotypic characteristics, to ensure patients at the greatest risk for severe illness receive appropriate care as early as possible (Liang et al., 2020). Wu et al. also used the Cox model to investigate the key risk factors and predicted the mortality rate of 21,392 COVID-19 patients based on demographic, clinical, and laboratory features and found that the mortality rate increased with time, especially for these critically ill patients (Wu et al., 2020).

Unfortunately, although the clinical features of COVID-19 patients have been reported in several recent publications (Gupta et al., 2020a; Xu et al., 2020), such as decreased lymphocytes and elevated CRP (Gupta et al., 2020a; Xu et al., 2020), the predictive powers and interpretations of these clinical features remain unclear. Additionally, since progression and outcome are critical for COVID-19 patients (Liang et al., 2020; Risch, 2020), timely monitoring from admission to outcome also has important clinical significance, making it possible to adjust treatment regimens in time, but this process is not entirely clear. Moreover, the foresight of a predictive model, as to how many days before discharge these features could accurately predict the patients' outcome, remains elusive. However, the association of these clinical features with phenotypic characteristics is also unclear. The robustness of the mortality prediction model along the timeline and the predictive power considering different preexisting diseases also need further exploration. Therefore, we performed this international, bi-institutional study to establish a mortality prediction model with the aim of early triaging and optimizing the resources.

METHODS

Ethical Approval

This study was approved by the Ethics Committee of Union Hospital, Tongji Medical College, Huazhong University of Science and Technology. Due to the retrospective nature of this study, the local institutional review board of the University of Würzburg waived the requirement for additional approval. This study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

Sample Description

Clinical data were collected from 1,441 COVID-19 patients from January 28, 2020, to March 29, 2020, at Wuhan Union Hospital (also called Wuhan cohort), China, for model development. Moreover, 96 patients with confirmed COVID-19 disease were collected from the University Hospital of Würzburg (also called Würzburg cohort), Germany, from March 6, 2020, to September 14, 2020, for independent test.

For the Wuhan cohort, more than 300 clinical features from hospital laboratory tests were recorded, and most patients have multiple sets of clinical features during their stay in the hospital. In addition, physical examinations, such as height, weight, temperature, sphygmus, systolic/diastolic pressure, respiratory rate, and heart rate, were performed upon admission of these COVID-19 patients. For robust analysis, clinical features that covered less than 30 samples, as well as samples containing fewer than three clinical features, were discarded (**Figure 1A**). After filtering out low-quality records, 1,352 patients and 130 clinical features were selected for systematic analysis. The average age of these patients was 58.22 (standard error: 14.90), and 50.52% of them were male, indicating a balanced gender. The minimal, maximal, and median duration from admission to discharge of the 1,352 patients is 0, 55, and 10 days, respectively. Among all of 1,352 COVID-19 patients, 1,221 patients survived and 131 died (**Supplementary Table S1**).

Clinical features (**Figure 1B**) from hospital laboratory tests were primarily composed of two parts: 101 numerical features, such as LDH and CRP, and 29 binary features, such as ABO blood type, Mp-IgM, and Mp-IgG. These clinical features were considered as candidate biomarkers for COVID19 mortality prognosis.

Phenotypic characteristics at admission (**Figure 1B**) were primarily composed of two parts: numerical and binary phenotypic characteristics. The numerical phenotypic characteristics included age, height, weight, temperature, sphygmus, systolic/diastolic pressure, respiratory rate, heart rate, and clinical classification. Binary phenotypic characteristics included records of gender, smoking status, and blood type.

Recent studies have already reported that the outcome of COVID-19 patients is greatly influenced by whether the
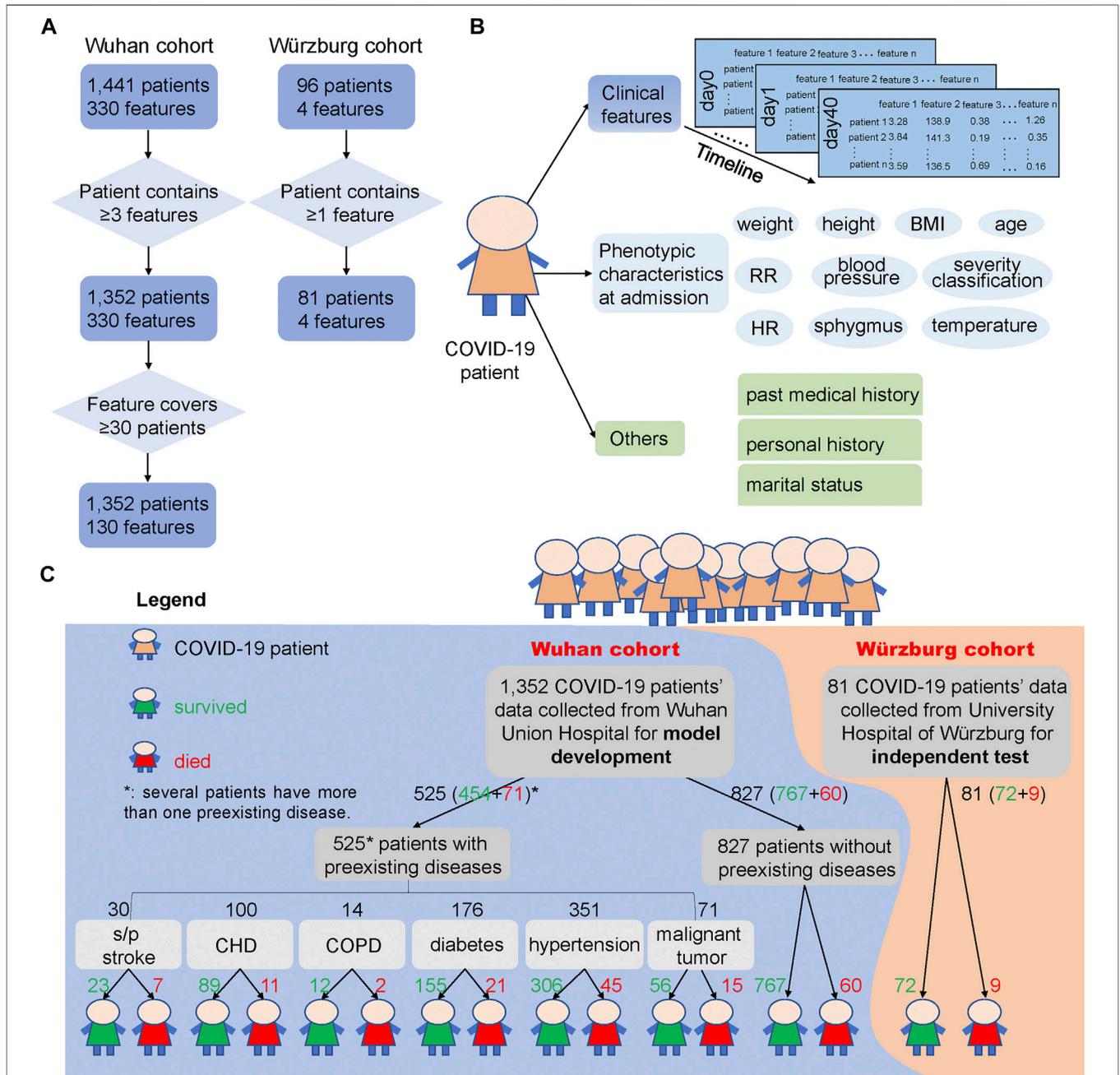
FIGURE 1 | COVID-19 patients and their clinical feature filtering process, phenotypic characteristics and clinical features used in this study, and the outcome of the two cohorts. **(A)** Process of filtering low-quality samples of the two cohorts. Here, 330 features in the Wuhan cohort were the union of 1,441 patients' clinical features and 130 features were the union of the 1,352 patients' clinical features after filtering. "≥3 features" says that the patients from the Wuhan cohort should contain at least three clinical features during the hospital stays, and "≥30 patients" says that the clinical features that collected from clinical laboratory should cover at least 30 patients and thus could be used for subsequent analysis. In the Würzburg cohort, "≥1 feature" says that the patient should contain at least one of the features from these four clinical features: lymphocyte (%), neutrophil count, LDH, and CRP. **(B)** Different types of clinical features and phenotypic characteristics used in the two cohorts. We used clinical features from hospital laboratory tests for developing the prediction model, and these clinical features were also used to test the association with phenotypic characteristics and other records. **(C)** Overview of samples used for model development and independent test. Samples of 1,352 COVID-19 patients from Wuhan Union Hospital (Wuhan cohort, the blue background) were used for building and testing the mortality prediction model, while samples of 81 COVID-19 patients from Germany (Würzburg cohort, the orange background) were used for independent test of the mortality prediction model. The green number represents the number of patients who survived from COVID-19, while the red number means the number of patients who died from COVID-19. Note that several patients have more than one preexisting disease.

patient has a preexisting disease (Azevedo et al., 2020; Zhou et al., 2020a; Williamson et al., 2020), such as CHD (Mai et al., 2020), hypertension (Itoh, 2020), and diabetes (Gupta et al., 2020b). Here, we divided 1,352 COVID-19 patients into seven groups: s/p stroke (23 survived and seven died), CHD (89, 11), chronic obstructive pulmonary disease (COPD) (12, 2), diabetes (155, 21), hypertension (306, 45), malignant tumor (56, 15), and those without preexisting diseases (767, 60) according to their past medical history (**Figure 1C**).

For the 96 patients in the Würzburg cohort, we have filtered out the patient who has not a single clinical feature among the four clinical features (lymphocyte (%), neutrophil count, LDH, and CRP) (**Figure 1A**). After this process, 81 samples were retained and utilized for independent test. For these 81 patients, their phenotypic characteristics including systolic pressure, diastolic pressure, temperature, heart rate, SpO_2, age, and respiratory rate were also used for analysis. The average age of these patients was 67.15 years (standard error: 15.17), which was significantly higher than that of patients in the Wuhan cohort (t-test, $p = 0.0005$). 62.96% of them are male, 53.67% of them have respiratory failure, and 41.46% of them need mechanical ventilation. Among them, 72 survived and nine died from COVID-19 (**Supplementary Table S2**).

Severity Classification

According to the diagnosis and treatment of pneumonia infected by the new novel coronavirus (the trial seventh edition) (National Health Commission of the People's Republic of China, 2020), the patient's severity classification was divided into three classifications, general, severe, and critical, according to their symptoms at admission. In this work, among 1,352 patients from the Wuhan cohort, 896 were in general, 393 were in severe, and 63 were in critical. For the Würzburg cohort, 24 were in general, 35 were in severe, and 22 were in critical. Here, we defined severity classification as follows: general as 1, severe as 2, and critical as 3.

Clinical Feature Profiling

Using patient samples at admission, all numerical clinical features were normalized to a range [0, 1]. These normalized data with an average abundance ≥0.001 were illustrated as boxplots using the R package "ggplot2". To illustrate differences between patients who survived and died, as well as between patients with or without preexisting diseases, principal coordinate analysis (PCoA) was performed using all patients' numerical clinical features at admission based on the Jaccard coefficient for distance measurement using the R package "vegan".

Feature Selection and Development of a Prediction Model Utilizing Clinical Features

To identify the most important clinical features that reflect differences among the samples, feature selection was employed for a deeper understanding of COVID-19 infection. We assessed the contribution of each clinical feature to facilitate the decisions of the algorithm. Considering both MeanDecreaseAccuracy and MeanDecreaseGini, the top five discriminatory clinical features were selected. Different Random Forest (RF) models were tested on the top five important clinical features, as well as their different combinations according to their importance.

To develop a mortality prediction model that is capable of distinguishing the outcome of COVID-19 patients, RF analysis was performed by randomForest() function in R (package "randomForest"). For the sample size larger than 100, we randomly selected 90% of samples as training set and 10% of samples as testing set using sample() function with replacement. In this process, replace parameter was set as true, which specifies using the Bootstrap method for random sampling. For each model, based on each training set, the important parameters ntree (number of decision trees contained in the RF model) and mtry (variable sampling values for each iteration) were trained and estimated with the out-of-bag (OOB) value. The importance was set as true for calculating the importance of each variable in the model, which was mainly used in conjunction with the importance() function. The proximity parameters were set as true for calculating the proximity matrix of the model, which is mainly used in conjunction with the MDSplot() function to realize the visualization of random forest. The na.action parameter specifies the methods for handling the missing values and was set as na.omi (that is, delete the samples with missing values of all features). Other parameters were set as default. A traversal search was performed on all clinical features to obtain the minimum OOB value. The value of mtry was determined by the OOB value (that is, the index of the minimum OOB value). Then, combining the outcome of COVID-19 patients, the mtry value was iterated to obtain an optimal ntree. This process was iterated 15,000 times or more to construct the most accurate model. When the error tree approaches stable, the minimum number of trees was the best value for ntree. This trained model was used for predicting the outcome of the testing set.

Evaluation of Prediction Models

To evaluate the performance of the RF model, we used several standard statistic parameters: accuracy, precision, sensitivity or recall, specificity, and F1 scores. Here, we defined the prediction result: survived-survived as TP and died-died as TN. The formulas of the parameters mentioned above are defined as follows:

$$accuracy = (TP + TN)/(TP + TN + FP + FN), \quad (1)$$

$$precision = TP/(TP + FP), \quad (2)$$

$$recall = TP/(TP + FN), \quad (3)$$

$$specificity = TN/(TN + FP), \quad (4)$$

$$F1 = 2* precision*recall/(precision + recall), \quad (5)$$

where TP, TN, FP, and FN stand for true-positive, true-negative, false-positive, and false-negative rates, respectively.

Correlation Analysis Between Phenotypic Characteristics and Clinical Features

To better understand the relationship between phenotypic characteristics and the mortality rate of patients, we used the Pearson coefficient to examine the correlation between

phenotypic characteristics and clinical features. Again, we organized these correlation values along the timeline to identify the dynamics of such correlations during treatment progression.

Evaluation of Prediction Models along the Timeline

Most patients have multiple sets of clinical features during their stay in the hospital, allowing for a series of mortality prediction models along the timeline. Here, we ordered these mortality prediction models in two directions: from admission forward to outcome to directly provide day-to-day guidance for clinics and from outcome backward to admission to evaluate the robustness and prediction power of the model against the time of hospital stay.

Development of High-Quality and Interpretable Binary Decision Tree for Clinical Diagnosis

Based on the five selected features, we aimed to develop a high-quality decision tree for clinical diagnosis. To train the RF model, the dataset was randomly separated into two groups: the training set (90% of entries) was applied to construct the mortality prediction model, and the testing set (10% of entries) was applied to validate the mortality prediction model. For datasets with a sample size of less than 100, we used 70% of the dataset for training and 30% for testing to reduce the contingency error. This process was iterated 15,000 times to construct the most accurate model. The most discriminative clinical feature was used as the root node of this binary decision tree, and the child nodes were hierarchically formed according to their distinguishing ability until all samples are completely distinguished. Finally, the decision tree was visualized by rpart() function in R (package "party").

Development of a Prediction Model for Different Preexisting Diseases

Considering the influence of preexisting diseases on the outcome in COVID-19 patients, we also used the first samples of patients with preexisting diseases as a training dataset to build the mortality prediction models: s/p stroke, CHD, COPD, diabetes, hypertension, malignant tumor, and those without preexisting diseases. For a dataset with a sample size larger than 100, we used 90% of the dataset for training and 10% for validation. For a dataset with a sample size smaller than 100, we used 70% of the dataset for training and 30% for testing to validate the model to reduce the contingency error.

Independent Test of the Mortality Prediction Model Using the Würzburg Cohort

To examine the reliability, interpretability, and foresight of our mortality prediction model developed based on the Wuhan cohort, 81 samples at admission from the Würzburg cohort were used for independent test. Pearson coefficient was also used to evaluate the association between the four clinical features (lymphocyte (%), neutrophil count, LDH, and CRP) and phenotypic characteristics (systolic pressure, diastolic pressure, temperature, heart rate, SpO_2, age, and respiratory rate).

RESULTS

In this study, we have recruited two independent cohorts from China (the Wuhan cohort) and Germany (the Würzburg cohort) for building and testing a mortality prediction model, respectively. The Wuhan cohort contained 1,352 COVID-19 patients from Wuhan Union Hospital, and it has been utilized for establishing a multi-feature and time-series aware machine learning models. The Würzburg cohort consists of 81 COVID-19 patients and has been used as an independent validation cohort.

Data Resource and General Profiles of COVID-19 Patients from Wuhan Cohort

1,352 patients were enrolled in the Wuhan cohort, who had more than three clinical features (such as neutrophil count, CRP, lymphocyte count, LDH, albumin, direct bilirubin, and creatine kinase) (Liang et al., 2020; Xiao et al., 2020) and detailed medicinal records from January 28, 2020, to March 29, 2020. The distribution of the number of patients with clinical laboratory tests on a daily basis, as well as the total number of diagnoses for each patient, is shown in **Supplementary Figure S1**. Among them, the mortality rates in patients with preexisting diseases: s/p stroke, coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD), diabetes, hypertension, and malignant tumor were 23.33, 11.00, 14.29, 11.93, 12.82, and 21.13%, respectively (**Supplementary Figures S2A,B**). These mortality rates were significantly higher (t-test, $p < 0.001$) than those in patients without preexisting diseases (mortality rate: 7.26%). PCoA showed that if we used all clinical features, these patients cannot be clearly separated (**Supplementary Figure S2C**). In addition, these patients could not be separated by whether they had a preexisting disease or not (**Supplementary Figures S2D–J**). This highlights the importance of clinical feature selection and developing the mortality prediction models to differentiate patients.

Development and Evaluation of Clinical Feature Selection and Mortality Prediction Model for Early Prognosis Based on Wuhan Cohort

We first developed a mortality prediction model based on patients' samples at admission, since such prediction is of paramount importance in clinics (Risch, 2020). This model took the clinical features and outcomes into consideration, aiming to optimize the medical resources, as well as preemptive therapy.

Before developing a mortality prediction model, we divided the 130 clinical features into two parts: 101 numerical features

For curves	Variables	Precision	Recall or Sensitivity	Specificity	F1 score	Accurary (%)	AUC(%)
Lymphocyte	Lymphocyte(%)	0.93	0.89	0.97	0.91	88.14±0.04	98.36
Neutrophil	Neutrophil	0.95	0.91	0.96	0.93	86.62±0.35	98.62
LDH	LDH	0.93	0.90	0.95	0.91	87.56±0.18	98.84
CRP	CRP	0.92	0.96	0.91	0.94	85.80±0.16	98.54
α-HBDH	α-HBDH	0.95	0.91	0.95	0.93	88.25±0.19	98.70
C2	LDH, α-HBDH	0.95	0.93	0.93	0.94	93.74±4.85	98.66
C3-1	Lymphocyte(%), LDH, α-HBDH	0.97	0.92	0.95	0.94	91.27±0.92	98.66
C3-2	Neutrophil, CRP, α-HBDH	0.90	0.94	0.94	0.92	89.68±0.14	98.83
C3-3	Neutrophil, LDH, α-HBDH	0.94	0.94	0.92	0.94	89.76±0.12	98.64
C3-4	Neutrophil, LDH, CRP	0.95	0.93	0.94	0.94	90.25±0.17	98.53
C3-5	Lymphocyte(%), neutrophil, LDH	0.96	0.94	0.91	0.95	91.91±0.13	98.54
C4-1	Lymphocyte(%), neutrophil, LDH, α-HBDH	0.98	0.94	0.93	0.96	91.03±0.11	98.52
C4-2	Lymphocyte(%), LDH, CRP, α-HBDH	0.97	0.93	0.93	0.95	90.89±0.13	98.71
C5	Lymphocyte(%), neutrophil, LDH, CRP, α-HBDH	0.97	0.91	0.96	0.94	91.42±0.17	99.83
C-all	All numerical features	0.96	0.93	0.94	0.95	91.72±0.20	98.84

FIGURE 2 | Clinical feature selection and mortality prediction results utilizing the first samples at admission. (A) Clinical feature selection based on 1,352 samples and 101 numerical features. Features were ranked by MeanDecreasedAccuracy and MeanDecreasedGini according to their importance. Considering both of these import RF parameters, we selected five important clinical features: lymphocyte (%), neutrophil count, LDH, CRP, and α-HBDH. (B) Comparison of receiver operating characteristic (ROC) curves and diagnostic performance of all numerical clinical features, subfeature combinations, as well as each selected single feature, using the first samples at admission (also referred to as admission-day 0). (C) The binary decision tree for predicting the outcome of COVID-19 patients based on the five selected clinical features at admission (admission-day 0). Annotations: Num: the number of patients in the predictor; T: the number of correctly matched patients; F: the number of mismatched patients. Here, the Num above the root node indicates the total samples used for building the binary decision tree.

and 29 binary features. For numerical features, those features which are identified with the average abundance ≥0.001 are shown in **Supplementary Figure S3**. 101 numerical clinical features with at least 30 samples' coverage were considered as the outcome predictors and were used to build the mortality prediction model. We used 90% of the samples for model training and 10% for testing to validate the model.

Combined MeanDecreaseAccuracy and MeanDecreaseGini (**Figure 2A**), lymphocyte (%), neutrophil count, C-reactive protein (CRP), lactic acid dehydrogenase (LDH), and α-hydroxybutyric dehydrogenase (α-HBDH) were selected for developing an optimized model, where lymphocyte (%) is an immune disorder indicator (Trowell, 1947), neutrophil count represents infection (Xie et al., 2020), CRP represents inflammatory response (Vermeire et al., 2004; Sabrina et al., 2012), and both LDH and α-HBDH represent tissue lesions (Sanwald and Kirk, 1966; Kishaba et al., 2014).

We then used these five selected numerical clinical features (lymphocyte (%), neutrophil count, CRP, LDH, and α-HBDH), as well as different combinations of the subset of these five clinical features according to their importance, for prediction (**Figure 2B**). Results showed that the performance of these five clinical features could be comparable to the results predicted by all numerical features. Considering the F1 score, accuracy, and AUC, the combination of lymphocyte (%), neutrophil count, and LDH also showed high performance, especially the performance of α-HBDH used alone (bold in **Figure 2B**). Several specified combinations of three out of these five clinical features, such as the combination of lymphocyte (%), neutrophil count, and LDH, also reached more than 91% accuracy and 99% AUC at admission. However, in clinics, these five features covered more types of clinical symptoms: lymphocyte (%) is an immune disorder indicator (Trowell, 1947), neutrophil count represents infection (Xie et al., 2020), CRP represents inflammatory response (Vermeire et al., 2004; Sabrina et al., 2012), and both LDH and α-HBDH represent tissue lesions (Sanwald and Kirk, 1966; Kishaba et al., 2014). Thus, we confirmed these five clinical features as credible biomarkers.

To benchmark with other classification algorithms, we also used FEAST (an expectation–maximization-based unsupervised learning method) (Shenhav et al., 2019) and JSD (Jensen–Shannon divergence) methods (Lin, 1991) to predict the outcome of COVID-19 patients based on all features, the top five features, and the top three features (**Supplementary Figure S4**). Results demonstrated that the RF model was more or equally credible for constructing the mortality prediction model. The neural network (Kriegeskorte and Golan, 2019) with two hidden layers (the first layer has 128 neurons and the second layer has eight neurons) also illustrated that RF model based on the combination of lymphocyte (%), neutrophil, LDH, CRP, and α-HBDH could best predict the outcome of COVID-19 patients (**Supplementary Figure S5**). Moreover, all three methods (RF, FEAST, JSD, and neural network) showed the best distinguishing power when using the top five clinical features to construct the model.

We also used the binary clinical features (such as urine occult blood, blood type, and COVID-19 nucleic acid) to build the

mortality prediction model (**Supplementary Figure S6A**). Based on the contribution of each feature, we selected urine protein (UPRO), urine occult blood (UOB), monospecific antibodies of blood type (Ab-monospecific-B), ABO blood type (ABO), and ketones (KET) for further model improvement. Their different combinations and performance are shown in **Supplementary Figure S6B**. Among them, the combination of UPRO, UOB, and KET (accuracy = 99.61%; AUC = 99.96%) was outstanding from the others, followed by UPRO, all binary features, and the combination of these five features.

Finally, we emphasized that all of the above results were based on the first samples at admission, since it is more important for the clinical prediction to utilize these samples. It was noticed that a recently published study used the final samples of COVID-19 patients for predicting their outcome (Yan et al., 2020), and we also used the final samples in the Wuhan cohort to assess our model based on five selected features (**Supplementary Figure S7**), with results showing high prediction accuracy. Yet, the prediction accuracy and AUC based on first samples at admission (**Figure 2B**) were comparable to those based on these final samples for the Wuhan cohort. These results confirmed again that patients with a high mortality rate could be accurately predicted at admission, which could be used for prioritizing critically ill patients to potentially reduce the mortality rate.

Clinical Features Have Profound Association with Phenotypic Characteristics in the Wuhan Cohort

Notable correlations were observed between phenotypic characteristics and clinical features associated with COVID-19 (**Figure 3** and **Supplementary Figures S8, S9**). Among 101 numerical clinical features, many of them have shown significant correlations with age, respiratory rate, and severity classification of patients. Expect for lymphocyte (%), neutrophil count, LDH, CRP, and α-HBDH were positively correlated with age ($p < 0.05$) along the timeline. Since the above analyses also confirmed that these five clinical features are tightly associated with patient outcomes (**Figure 2**), these associations partially verified the fact that elder patients were more likely to die from COVID-19. LDH, CRP, and α-HBDH were also positively correlated with respiratory rate and severity classification ($p < 0.05$) in patients (896 were in general, 392 were in severe, and 63 were in critical), illustrating the importance of these phenotypic characteristics on outcome in COVID-19 patients. The result also showed dynamic changes in the associations of these clinical features with phenotypic characteristics over time, especially for the five clinical features used for model prediction.

Time-Series Analysis Reveals That the Mortality Prediction Model Is Very Robust along the Timeline

Evaluation of the mortality prediction model along the timeline forward from admission day as the start point: Because these clinical features are dynamic along the timeline, and in clinics, the

FIGURE 3 | Associations between clinical features and phenotypic characteristics based on several representative time points. **(A)** Based on admission day (admission-day 0). **(B)** Based on the third day after admission (admission-day 1). **(C)** Based on the sixth day after admission (admission-day 2). **(D)** Based on the ninth day after admission (admission-day 3). **(E)** Based on the admission day (admission-day 4). **(F)** Based on the third day after admission (admission-day 5). **(G)** Based on the sixth day after admission (admission-day 6). **(H)** Based on the ninth day after admission (admission-day 9). Note: *represents a significant correlation between a phenotypic characteristic and a clinical feature (Pearson correlation: $p < 0.05$).

progression and outcome of patients are critical (Liang et al., 2020; Risch, 2020). Therefore, we used the admission day of each patient as the start point and built mortality prediction models day by day after admission along the timeline. The number of samples enrolled on a daily basis is shown in **Figure 4A** from admission-day 0. We used 90% of the dataset for training and 10% for testing. For datasets with a sample size of less than 100, we used 70% of the dataset for training and 30% as a test set for validation. Since the sample number was less than 50 for patients who stayed in the hospital longer than 40 days, we only used the dataset from admission-day 0 to admission-day 40 to build the time-series mortality prediction models. Results confirmed that our mortality prediction model was very robust over time,

suggesting that according to the prediction outcome of patients, clinics could adjust the treatment plan at any time, which could provide higher quality treatment for patients.

Evaluation of the mortality prediction model along the timeline backward from discharge day as the start point: To prove the robustness of our mortality prediction model and how many days in advance it could predict the outcome of COVID-19 patients, we used the discharge day of each patient as the start point. Prediction accuracies were evaluated backward day by day (**Figure 4B**) from discharge-day 0. The mortality prediction model based on five clinical features also reached more than 91% accuracy and 99% AUC (usually 10 days or more in advance of the outcome) (**Figure 4B**), confirming this mortality prediction

FIGURE 4 | Evaluation of prediction results based on time series with five selected features. The prediction results were evaluated based on time series using the admission-day 0 **(A)** and the discharge-day 0 **(B)** as start points. The first *y*-axis represents the number of samples enrolled, while the second *y*-axis represents the estimation scores. The purple bar represents the number of patients discharged from the hospital on that day, while the red bar indicates the number of patients admitted to the hospital on that day.

model is very robust over time when patients were in the hospital and indicating the strong association of these five clinical features with the progression of treatment.

The Highly Accurate and Interpretable Binary Decision Tree for Clinical Diagnosis

To make the prediction interpretable, we also generated a series of decision trees (along the timeline) for assisting clinical diagnosis based on the Wuhan cohort. The decision tree is hierarchically organized by the distinguishing ability of these five clinical features based on the first samples at admission (**Figure 2C**). LDH could distinguish 87% of samples with more than 96% accuracy and was used as the root node of this decision tree. The remaining 13% of samples were differentiated by a combination of these five clinical features. The binary decision tree of the final samples at discharge was simpler than that of the first samples at admission (**Supplementary Figure S10E**). The decision trees based on other time points are shown in **Supplementary Figures S10A–D**, confirming that using these five clinical features was more comprehensive and precise. These results

also suggested that the mortality prediction model based on the admission samples, rather than the discharge samples, could already provide outcome prediction and clinical guidance for personalized treatment with high fidelity.

The binary decision tree, either based on samples at admission or based on discharge, was also highly interpretable for clinical diagnosis. The elevated LDH was associated with patients' death: LDH larger than 445 u/l was a significant risk factor related to death in cases with severe COVID-19 (Zhou et al., 2020b; Li et al., 2020), which was consistent with our results. The increased level of α-HBDH was also found as a critical risk factor associated with the severity of COVID-19 patients (Dong et al., 2020). The decreased amount of lymphocyte (lymphopenia) and neutrophil (neutrophilia), together with the increased number of CRP and LDH, showed the immunological response to the virus, followed by severe virus infection (Frater et al., 2020; Lippi and Plebani, 2020). In summary, current published clinical evidence could well support our decision tree.

Prediction Power Considering Different Preexisting Diseases

For different preexisting diseases, the clinical features that can accurately mark the COVID-19 patients' outcomes are generally different. Previous studies have shown that preexisting disease increases the risk of COVID-19 mortality rate (Williamson et al., 2020). We also used the six preexisting diseases to evaluate the mortality prediction model based on the five selected clinical features (**Supplementary Figure S11**).

Out of the five selected clinical features, feature combinations should be different for each of the different preexisting diseases. Therefore, for each of the preexisting diseases, we performed feature importance evaluation before mortality prediction model evaluation. For patients with s/p stroke (**Supplementary Figure S11A**), considering F1 score, accuracy, and AUC, the combination of LDH, CRP, and α-HBDH showed the highest performance, followed by the combination of LDH and α-HBDH, then all five features. The results for patients with CHD are illustrated in **Supplementary Figure S11B**. Except for using the five features, the combination of LDH, CRP, and α-HBDH showed the highest performance. For patients with COPD, a combination of neutrophil count, lymphocyte (%), and LDH showed the highest performance (**Supplementary Figure S11C**). For patients with diabetes, among all combinations of clinical features, LDH showed the highest performance (**Supplementary Figure S11D**), indicating that LDH could be used to distinguish the outcome of COVID-19 patients. For patients with hypertension, results indicated that a combination of neutrophil count, lymphocyte (%), and LDH could be used as biomarkers for predicting the outcome of COVID-19 patients with hypertension (**Supplementary Figure S11E**). For patients with malignant tumor, the combination of all five features showed the highest performance, followed by the combination of neutrophil count and lymphocyte (%) (**Supplementary Figure S11F**). For patients without preexisting diseases, results showed that we can use lymphocyte (%), LDH, CRP, and α-HBDH to accurately predict the outcome of these patients (**Supplementary Figure S11G**).

FIGURE 5 | Independent test results on the Würzburg cohort utilizing the first samples at admission. **(A)** The number of patients categorized by their duration of the hospital stay. **(B)** The associations between clinical features (LDH, CRP, lymphocyte (%), and neutrophil count) and phenotypic characteristics. **(C)** Comparison of receiver operating characteristic (ROC) curves and diagnostic performance of four clinical features, the combinations of a subset of features, as well as each selected single feature, using the first samples at admission. Note: * represents a significant correlation between a phenotypic characteristic and a clinical feature (Pearson correlation: $p < 0.05$).

Evaluation of the Mortality Prediction Model Using the Independent Würzburg Cohort

The reliability, interpretability, and foresight of our mortality prediction model were further confirmed in another independent cohort collected from Germany, the Würzburg cohort, with samples collected from March to September 2020 (**Figure 5**). For the patients in the Würzburg cohort, their duration of stay in the hospital is usually 5–20 days (**Figure 5A**). All samples used in the Würzburg cohort were the patient samples at admission.

We used four clinical features (lymphocyte (%), neutrophil count, LDH, and CRP), as well as their different combinations to test our

mortality prediction model (**Figure 5C**). Considering F1 score, accuracy, and AUC, the combination of LDH, lymphocyte (%), neutrophil count, and CRP (accuracy = 97.33%; AUC = 85.71%) showed the highest performance among different combinations. Other combinations, such as the combination of LDH, lymphocyte (%), and neutrophil count (accuracy = 93.67%, AUC = 87.83%) and the combination of LDH and lymphocyte (%) (accuracy = 94.81%, AUC = 84.25%), also performed well. When only one clinical feature was used, LDH (accuracy = 91.30%, AUC = 85.56%) showed the highest performance, which was consistent with the results on the Wuhan cohort and a previous study (Liang et al., 2020).

From the Pearson correlation analysis (**Figure 5B**) between these four clinical features (LDH, lymphocyte (%), neutrophil count, and CRP) and the phenotypic characteristics (systolic pressure, diastolic pressure, temperature, heart rate, SpO_2, age, and respiratory rate), we could observe that there was a significantly negative correlation between lymphocyte (%) and age ($p < 0.05$), which was consistent with the general pattern of COVID-19 patients. CRP was significantly positively correlated with temperature ($p < 0.05$), which was consistent with the result in the Wuhan cohort.

Furthermore, as the duration of stay in the hospital of patients is usually 5–20 days, the strong foresight of the mortality prediction model has again been validated on the Würzburg cohort. Furthermore, one male patient aged 54 has a hospital stay of 93 days before recovery, and our mortality prediction model has successfully predicted his outcome.

DISCUSSIONS AND CONCLUSION

Our study enrolled two independent cohorts of COVID-19 patients for reliable, interpretable, and universal mortality model evaluation. Through multiple analyses including RF analysis, association analysis, time-series analysis, etc., the mortality prediction model was established, evaluated, and achieved clinically creditable prediction power on the Wuhan cohort and Würzburg cohort.

The mortality prediction model proposed in this study could help identify critically ill patients early and provide preferential treatment for each individual. Firstly, the five important clinical features (lymphocyte (%), neutrophil count, CRP, LDH, and α-HBDH) were identified. These five features could reflect several important aspects of disease development, such as viral infection (Trowell, 1947), coexistence of other infections (Xie et al., 2020), immune reaction during pneumonia (Sabrina et al., 2012), the severity of inflammatory response (Vermeire et al., 2004), tissue/cell damage, and cardiac injury (Sanwald and Kirk, 1966; Kishaba et al., 2014), which could provide more information to monitor the progression of patients. Secondly, these five features could be used for predicting the outcome of COVID-19 patients with high accuracy. Thirdly, the foresight of the mortality prediction model was strong up to as early as 40 days or more before discharge. This indicates that our model could allow resource optimization to be conducted many days ahead, and physicians can make a preliminary judgment on the prognosis of patients according to this model to prompt the choice of clinical intervention in later stages.

Our mortality prediction model shows superior prediction power at different time points during the course of the disease. Robust prediction power at different time points (**Figure 4**) also suggests that the mortality prediction model provides important indicators for disease monitoring, indicating early clinical intervention for clinical treatment. Our mortality prediction model also shows superior prediction power for different preexisting diseases of patients, indicating the robustness of the mortality prediction model. These results could serve well as the basis for personalized treatment of COVID-19 patients.

Our finding in the Wuhan cohort (model development) has also been tested in an independent cohort from Germany (Würzburg cohort). Although the international aspects such as the ethnicities, healthcare systems, hygienic measures, local regulations, and management strategies, as well as their average age (t-test, $p = 0.0005$), are different in these two cohorts, our mortality prediction model has also shown the high prediction power in tens of days ahead of patients' discharge, underlining the robustness and the foresight of this model.

The second COVID-19 wave in Europe is ongoing. This mortality prediction model has been validated at a European center and might provide a useful instrument for triaging the patients and optimizing the resources. Because we have a series of mortality prediction models with constant high accuracy along with the whole duration of patients' stay in the hospital, we could adjust treatment for possibly serious patients on a day-to-day basis to reduce the mortality rate of patients with COVID-19 as much as possible. In addition, our study also provides new insight into the mortality prediction model's application value in other infectious disease outbreaks in the future.

In conclusion, this study has established a mortality prediction model that allowed early classification of COVID-19 towards personalized treatment in these patients, not only at admission but also along the treatment timeline. This model may represent a valuable tool for triaging and optimizing the resources in patients with COVID-19 infection worldwide.

DATA AVAILABILITY STATEMENT

The raw data used in this study are not publicly available due to the confidential policy of National Health Commission of China, as well as the General Data Protection Regulation (GDPR) of European Union (EU), but are available from the corresponding author Xiaohua Hou upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, and University of Würzburg.

AUTHOR CONTRIBUTIONS

XH and KN designed and conceived the study. TB, DW, LZ, JS, XZ, YJ, DG, NI, KK, LR, and HE collected the data. KN, XZ, and PY developed the model. KN, XZ, PY, YZ, and HC analyzed the data. KN, TB, XZ, and PY wrote the manuscript. XH, KN, TB, XZ, PY, XZ, HC, QY, LR, and HE revised the manuscript. All authors have read and approved the final manuscript.

FUNDING

REFERENCES

Ammar, M. A., Sacha, G. L., Welch, S. C., Bass, S. N., Kane-Gill, S. L., Duggal, A., et al. (2021). Sedation, Analgesia, and Paralysis in COVID-19 Patients in the Setting of Drug Shortages. *J. Intensive Care Med.* 36 (2), 157–174. doi:10.1177/0885066620951426

Azevedo, R. B., Botelho, B. G., Hollanda, J. V. G., Ferreira, L. V. L., Junqueira de Andrade, L. Z., Oei, S., et al. (2020). Covid-19 and the Cardiovascular System: a Comprehensive Review. *J. Hum. Hypertens.* doi:10.36660/ijcs.20200150

Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA: Association for Computing Machinery, 785–794.

Dong, Y., Zhou, H., Li, M., Zhang, Z., Guo, W., Yu, T., et al. (2020). A Novel Simple Scoring Model for Predicting Severity of Patients With SARS-CoV-2 Infection. *Transboundary emerging Dis.* 67 (6), 2823–2829. doi:10.1111/tbed.13651

Epidemiology Working Group for Ncip Epidemic Response CCfDC, Prevention (2020). The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi.* 41 (2), 145–151. doi:10.3760/cma.j.issn.0254-6450.2020.02.003

Frater, J. L., Zini, G., d'Onofrio, G., and Rogers, H. J. (2020). COVID-19 and the Clinical Hematology Laboratory. *Int. J. Lab. Hematol.* 42 (Suppl. 1), 11–18. doi:10.1111/ijlh.13229

Gomes, D. C. d. S., and Serra, G. L. d. O. (2021). Machine Learning Model for Computational Tracking and Forecasting the COVID-19 Dynamic Propagation. *IEEE J. Biomed. Health Inform.* 25 (3), 615–622. doi:10.1109/jbhi.2021.3052134

Gupta, S., Hayek, S. S., Wang, W., Chan, L., Mathews, K. S., Melamed, M. L., et al. (2020a). Factors Associated With Death in Critically Ill Patients With Coronavirus Disease 2019 in the US. *JAMA Intern. Med.* 180 (11), 1436–1447. doi:10.1001/jamainternmed.2020.3596

Gupta, R., Hussain, A., and Misra, A. (2020b). Diabetes and COVID-19: Evidence, Current Status and Unanswered Research Questions. *Eur. J. Clin. Nutr.* 74 (6), 864–870. doi:10.1038/s41430-020-0652-1

Ikemura, K., Bellin, E., Yagi, Y., Billett, H., Saada, M., Simone, K., et al. (2021). Using Automated Machine Learning to Predict the Mortality of Patients with COVID-19: Prediction Model Development Study. *J. Med. Internet Res.* 23 (2), e23458. doi:10.2196/23458

Itoh, H. (2020). A New normal for Hypertension Medicine with Coronavirus Disease-2019 (COVID-19): Proposal from the President of the Japanese Society of Hypertension. *Hypertens. Res.* 43 (9), 857–858. doi:10.1038/s41440-020-0497-y

Kishaba, T., Tamaki, H., Shimaoka, Y., Fukuyama, H., and Yamashiro, S. (2014). Staging of Acute Exacerbation in Patients With Idiopathic Pulmonary Fibrosis. *Lung.* 192 (1), 141–149. doi:10.1007/s00408-013-9530-0

Kriegeskorte, N., and Golan, T. (2019). Neural Network Models and Deep Learning. *Curr. Biol.* 29 (7), R231–r236. doi:10.1016/j.cub.2019.02.034

Li, X., Xu, S., Yu, M., Wang, K., Tao, Y., Zhou, Y., et al. (2020). Risk Factors for Severity and Mortality in Adult COVID-19 Inpatients in Wuhan. *J. Allergy Clin. Immunol.* 146 (1), 110–118. doi:10.1016/j.jaci.2020.04.006

Liang, W., Yao, J., Chen, A., Lv, Q., Zanin, M., Liu, J., et al. (2020). Early Triage of Critically Ill COVID-19 Patients Using Deep Learning. *Nat. Commun.* 11 (1), 3543. doi:10.1038/s41467-020-17280-8

Lin, J. (1991). Divergence Measures Based on the Shannon Entropy. *IEEE Trans. Inform. Theor.* 37 (1), 145–151. doi:10.1109/18.61115

Lippi, G., and Plebani, M. (2020). The Critical Role of Laboratory Medicine During Coronavirus Disease 2019 (COVID-19) and Other Viral Outbreaks. *Clin. Chem. Lab. Med.* 58 (7), 1063–1069. doi:10.1515/cclm-2020-0240

SUPPLEMENTARY MATERIAL

Mai, F., Del Pinto, R., and Ferri, C. (2020). COVID-19 and Cardiovascular Diseases. *J. Cardiol.* 76 (5), 453–458. doi:10.1016/j.jjcc.2020.07.013

National Health Commission of the People's Republic of China (2020). Diagnosis and Treatment of Pneumonia Infected by the New Novel Coronavirus (The Trial Fifth Edition) Medical Letter From the National Health Office.

Rasmussen, S., Sperling, P., Poulsen, M. S., Emmersen, J., and Andersen, S. (2020). Medical Students for Health-Care Staff Shortages during the COVID-19 Pandemic. *The Lancet.* 395 (10234), e79–e80. doi:10.1016/s0140-6736(20)30923-5

Risch, H. A. (2020). Early Outpatient Treatment of Symptomatic, High-Risk Covid-19 Patients that Should Be Ramped-Up Immediately as Key to the Pandemic Crisis. *Am. J. Epidemiol.* 189 (11), 1218–1226. doi:10.1093/aje/kwaa093

Sabrina, P., Pierre-Frederic, K., and Nicolas, V. (2012). CRP Pro-inflammatory Signalling in Atherosclerosis: Myth or Reality? *Curr. Signal Transduction Ther.* 7 (2), 142–148. doi:10.2174/157436212800376681

Salyer, S. J., Maeda, J., Sembuche, S., Kebede, Y., Tshangela, A., Moussif, M., et al. (2021). The First and Second Waves of the COVID-19 Pandemic in Africa: a Cross-Sectional Study. *The Lancet.* 397 (10281), 1265–1275. doi:10.1016/s0140-6736(21)00632-2

Sanwald, R., and Kirk, J. E. (1966). α-Hydroxybutyric Dehydrogenase Activity of Human Vascular Tissue. *Nature.* 209 (5026), 912–913. doi:10.1038/209912a0

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Pallavicini, A., Ruscio, M., et al. (2020). Searching for SARS-COV-2 on Particulate Matter: A Possible Early Indicator of COVID-19 Epidemic Recurrence. *Int. J. Environ. Res. Public Health.* 17 (9), 2986. doi:10.3390/ijerph17092986

Shenhav, L., Thompson, M., Joseph, T. A., Briscoe, L., Furman, O., Bogumil, D., et al. (2019). FEAST: Fast Expectation-Maximization for Microbial Source Tracking. *Nat. Methods.* 16 (7), 627–632. doi:10.1038/s41592-019-0431-x

Sirleaf, E. J., and Clark, H. (2021). Report of the Independent Panel for Pandemic Preparedness and Response: Making COVID-19 the Last Pandemic. *The Lancet.* 398 (10295), 101–103. doi:10.1016/s0140-6736(21)01095-3

Trowell, O. A. (1947). Function of the Lymphocyte. *Nature.* 160 (4076), 845–846. doi:10.1038/160845a0

Vermeire, S., Van Assche, G., and Rutgeerts, P. (2004). C-Reactive Protein as a Marker for Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* 10 (5), 661–665. doi:10.1097/00054725-200409000-00026

Wahlster, S., Sharma, M., Lewis, A. K., Patel, P. V., Hartog, C. S., Jannotta, G., et al. (2021). The Coronavirus Disease 2019 Pandemic's Effect on Critical Care Resources and Health-Care Providers. *Chest.* 159 (2), 619–633. doi:10.1016/j.chest.2020.09.070

Wang, H., Wang, L., Lee, E. H., Zheng, J., Zhang, W., Halabi, S., et al. (2021). Decoding COVID-19 Pneumonia: Comparison of Deep Learning and Radiomics CT Image Signatures. *Eur. J. Nucl. Med. Mol. Imaging.* 48 (5), 1478–1486. doi:10.1007/s00259-020-05075-4

Watson, S. I., and Lilford, R. J. (2021). Global COVID-19 Vaccine Roll-Out: Time to Randomise Vaccine Allocation?. *The Lancet.* 397 (10287), 1804–1805. doi:10.1016/s0140-6736(21)00895-3

Williamson, E. J., Walker, A. J., Bhaskaran, K., Bacon, S., Bates, C., Morton, C. E., et al. (2020). Factors Associated with COVID-19-Related Death Using OpenSAFELY. *Nature.* 584 (7821), 430–436. doi:10.1038/s41586-020-2521-4

Wu, R., Ai, S., Cai, J., Zhang, S., Qian, Z., Zhang, Y., et al. (2020). Predictive Model and Risk Factors for Case Fatality of COVID-19: A Cohort of 21,392 Cases in Hubei, China. *The Innovation.* 1 (2), 100022. doi:10.1016/j.xinn.2020.100022

Xiao, L.-s., Zhang, W.-F., Gong, M.-c., Zhang, Y.-p., Chen, L.-y., Zhu, H.-b., et al. (2020). Development and Validation of the HNC-LL Score for Predicting the Severity of Coronavirus Disease 2019. *EBioMedicine.* 57, 102880. doi:10.1016/j.ebiom.2020.102880

Xie, X., Shi, Q., Wu, P., Zhang, X., Kambara, H., Su, J., et al. (2020). Single-Cell Transcriptome Profiling Reveals Neutrophil Heterogeneity in Homeostasis and Infection. *Nat. Immunol.* 21 (9), 1119–1133. doi:10.1038/s41590-020-0736-z

Xu, X.-W., Wu, X.-X., Jiang, X.-G., Xu, K.-J., Ying, L.-J., Ma, C.-L., et al. (2020). Clinical Findings in a Group of Patients Infected With the 2019 Novel Coronavirus (SARS-Cov-2) Outside of Wuhan, China: Retrospective Case Series. *Bmj.* 368, m606. doi:10.1136/bmj.m606

Yan, L., Zhang, H.-T., Goncalves, J., Xiao, Y., Wang, M., Guo, Y., et al. (2020). An Interpretable Mortality Prediction Model for COVID-19 Patients. *Nat. Mach Intell.* 2 (5), 283–288. doi:10.1038/s42256-020-0180-7

Zhou, X., Bai, T., Meckel, K., Song, J., Jin, Y., Kortüm, K. M., et al. (2020a). COVID-19 Infection in Patients with Multiple Myeloma: a German-Chinese Experience from Würzburg and Wuhan. *Ann. Hematol.* 100 (3), 843–846. doi:10.1007/s00277-020-04184-2

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., et al. (2020b). Clinical Course and Risk Factors for Mortality of Adult Inpatients with COVID-19 in Wuhan, China: a Retrospective Cohort Study. *The Lancet.* 395 (10229), 1054–1062. doi:10.1016/s0140-6736(20)30566-3

Zhu, W., Xie, L., Han, J., and Guo, X. (2020). The Application of Deep Learning in Cancer Prognosis Prediction. *Cancers.* 12 (3). doi:10.3390/cancers12030603